

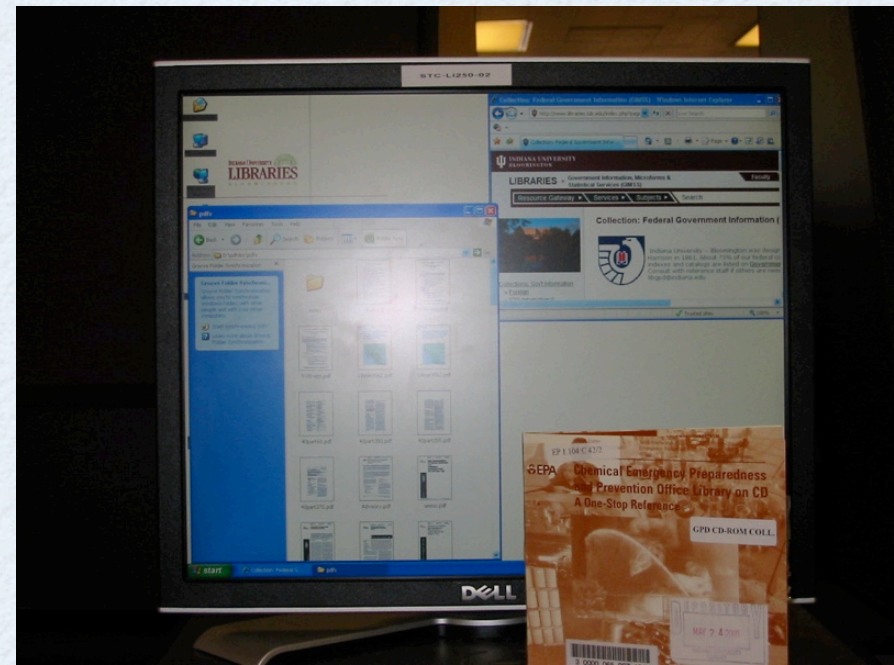
Sustainable Web Access to the SUDOC CD-ROMs:  
Experiments in the Preservation of Digital Documents  
in Legacy Formats

Geoffrey Brown  
Indiana University  
Department of Computer Science









Indiana University Department of

ComputerScience

in

IU Home|I

Iso

Edit | U

Search Site

Go

Home Page

Contacts

Database

Search

Help

RSS

Title

Chemical Emergency Preparedness and Prevention Office library on CD  
[computer file] :  
a one-stop reference

Call number

EP 1.104: C 42/2

System Requirements

76.8 MB hard drive space  
Adobe Acrobat Reader + Search version 3.0 (included on this CD-ROM)  
Microsoft Windows 95 or Windows 98  
Adobe Acrobat for Macintosh.

Installation Notes

Browse CD Images

[http://www.cs.indiana.edu/sudoc/image\\_30000065997136/](http://www.cs.indiana.edu/sudoc/image_30000065997136/)  
VMware

Iso





# Index of /sudoc/image\_30000065997136/30000065997136

[Name](#)

[Last modified](#)

[Size](#)

[Description](#)



[Parent Directory](#)



[pdfdoc/](#)

Apache/2.0.52 (Red Hat)

Single Disk Document  
30000065997136

Index of /

Name	Last Modified	Size
------	---------------	------

Upload  Browse...

Applet vnc started

SVP 3

Recycle Bin

VLC media player

SUDOC Share

Adobe Reader 3

Mozilla Firefox

PDFCreator

ssa3d30.ocx

000308\_1153 (D:)

File Edit View Favorites Tools Help

Back Forward Search Folders

Address D:\

Go

File and Folder Tasks

- Publish this folder to the Web
- Share this folder

Other Places

- My Computer
- My Documents
- Shared Documents
- My Network Places

Details

pdfdoc

start 000308\_1153 (D:) 9:49 AM



# ISSUES

- Collection isn't readily accessed
  - distributed across 1450 libraries
  - not always cataloged
  - requires physical access
- Collection not easily searched
  - minimal metadata
  - obsolete file formats
- ...



# ISSUES CONT.

- Software becoming obsolete
  - lotus 1-2-3 no longer supported in Office
  - many titles have proprietary software
- Knowledge of software use eroding
  - many MSDOS, windows 3.1 applications
  - install instructions often incorrect for later versions of windows



# FDLP GUIDELINES

1. All depository libraries should have a written policy regarding public services for Government information in electronic formats. This policy should contain provisions for no-fee access to computer workstations with CD-ROMs, diskettes, and the Internet. These should be equal to or exceed the services provided for other collections of the library.

2. All depository libraries should offer access to Government electronic information disseminated through the Federal Depository Library Program to the general public. Attempts must be made to obtain hardware that meets the latest "Recommended Specifications for Public Access Work Stations in Federal Depository Libraries" as published annually in Administrative Notes.

3. All depository libraries should make tangible electronic products and services (CD-ROMs, floppy diskettes) which they select available to the general public in a timely manner. For example, if a product is not currently loaded and/or supported on a depository library's computer workstation, the depository library should attempt to provide access to it within a designated time frame as determined by each library. If the depository library is unable to provide adequate access to and technical support of tangible electronic products, circulation of those products should be made available in accordance with the library's circulation policies of other non-governmental tangible electronic products or other depository resources



# OUR STRATEGY

- Provide web access and search for collection
- Provide migration to web formats
- Preserve access to collections through web based emulation
- Provide mechanisms for sparse community to contribute usage metadata



# REMAINDER OF TALK

- Building a “Raw” Archive
- Improving Browsing Through Migration
- Emulation

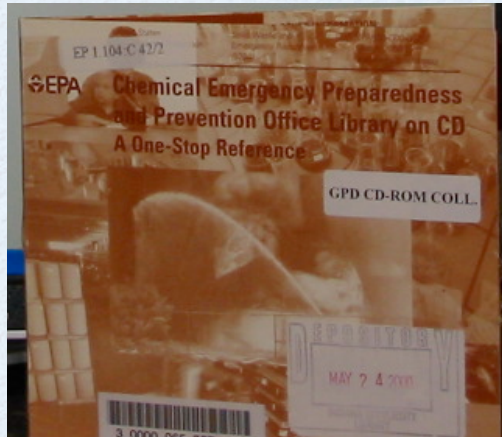


# RAW ARCHIVE

- Cheap to build (5 minutes / CD), few TBytes storage
- Easy to enable browsing of “raw” files
- Issues (lots)
  - obsolete files, programs
  - limited search capability
  - copyright / access control issues
  - limited metadata
  - no way to validate bits
  - no universal “id” for items in collection
  - IU collection incomplete



# BUILDING THE "RAW" ARCHIVE



Chemical Emergency Preparedness ...  
EP 1.104: C 42/2  
30000065997136

Work-study student using  
isorecorder on PC

3000065997136.iso

<http://isorecorder.alexfeinman.com/isorecorder.htm>



# "ISO" FILES

An ISO image (.iso) is a disk image of an ISO 9660 file system. ISO 9660 is an international standard originally devised for storing data on CD-ROM. More loosely, it refers to any optical disc image, even a UDF image.

As is typical for disc images, in addition to the data files that are contained in the ISO image, it also contains all the filesystem metadata, including boot code, structures, and attributes. All of this information is contained in a single file. These properties make it an attractive alternative to physical media for the distribution of software that requires this additional information as it is simple to retrieve over the Internet

[http://en.wikipedia.org/wiki/Iso\\_file](http://en.wikipedia.org/wiki/Iso_file)



# BUILDING DATABASE

30000065997136

Chemical Emergency Preparedness ...  
EP 1.104: C 42/2  
30000065997136

Library staff  
lookup "key"  
2886148

Workstudy Student



Database Entry

The diagram illustrates the process of building a database entry. Two distinct paths, represented by blue arrows, converge on a central point labeled 'Database Entry'. The left path originates from a numerical identifier '30000065997136' and a library staff lookup key '2886148'. The right path originates from a text description 'Chemical Emergency Preparedness ... EP 1.104: C 42/2' and the same numerical identifier '30000065997136'. The text 'Workstudy Student' is positioned near the right path. The entire diagram is set against a light blue background with a darker blue header containing the title 'BUILDING DATABASE'.



# CATALOG RECORD

2886148

script to z39.50 server

marcxml

xml to html

***Chemical Emergency Preparedness and Prevention Office library on CD [computer file] : a one-stop reference***

Washington, DC : U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, 2000.

1 computer optical disc ; 4 3/4 in.

76.8 MB hard drive space; Adobe Acrobat Reader + Search version 3.0 (included on this CD-ROM); Microsoft Windows 95 or Windows 98; Adobe Acrobat for Macintosh. "Issued April 2000."

"EPA 550-C00-001"--disc label.

Contains information on chemical accident prevention and emergency preparedness and response.

**1.** Chemicals--Safety measures. **2.** Hazardous substances--Safety measures.

1 37

#9138BB

[Full Record](#)



# ACCESSING FILES ON ISO

- ISO files mounted using iso9660 file system on Linux
- We mount on demand using “automounter”
- Export all iso file systems through Apache
- Web interface to simplify browsing of iso file systems










# "FREE" INTERFACE

Index of /sudoc/image\_30000042367908/30000042367908/

http://www.cs.indiana.edu/sudoc/image\_30000042367908/30000042367908/

nytimes slashdot IUB Libraries Google Reader sudoc/ B649 C335

## Index of /sudoc/image\_30000042367908/30000042367908

Name	Last modified	Size	Description
 <a href="#">Parent Directory</a>		-	
 <a href="#">adobemac/</a>	05-Oct-1995 06:35	-	
 <a href="#">adobewin/</a>	05-Oct-1995 06:35	-	
 <a href="#">agr4.pdx</a>	05-Oct-1995 06:35	965	
 <a href="#">agr4/</a>	05-Oct-1995 06:35	-	
 <a href="#">agr4all.pdf</a>	05-Oct-1995 06:35	4.7M	
 <a href="#">readme.txt</a>	05-Oct-1995 06:35	6.5K	

Apache/2.0.52 (Red Hat) Server at troy.cs.indiana.edu Port 80



# AN IMMEDIATE BENEFIT

[Advanced Search](#)  
[Preferences](#)[New! View and manage your web history](#)

Web

Results 1 - 10 of about 36,000 from [www.cs.indiana.edu/sudoc](http://www.cs.indiana.edu/sudoc). (0.17 seconds)

## [Index of /sudoc](#)

[DIR] Parent Directory - [DIR] image\_30000026394985/ 07-Sep-2007 05:00 - [DIR] image\_30000026404321/ 07-Sep-2007 05:00 - [DIR] ...

[www.cs.indiana.edu/sudoc/](http://www.cs.indiana.edu/sudoc/) - 210k - [Cached](#) - [Similar pages](#) - [Note this](#)

## [Index of /sudoc/image\\_30000065775821](#)

[DIR] Parent Directory - [DIR] 30000065775821\_1/ 31-Dec-1969 19:00 - [DIR] 30000065775821\_2/ 18-Feb-1998 10:06 - [DIR] 30000065775821\_3/ 24-Feb-1998 09:46 ...

[www.cs.indiana.edu/sudoc/image\\_30000065775821/](http://www.cs.indiana.edu/sudoc/image_30000065775821/) - 2k - [Cached](#) - [Similar pages](#) - [Note this](#)

## [\[PDF\] Contents](#)

File Format: PDF/Adobe Acrobat

Table of contents << page 1 >> << page 2 >> << page 3 >> << page 4 >> << page 5 >>. v.

Table of contents << page 1 >> << page 2 >> << page 3 >> << page 4 ...

[www.cs.indiana.edu/sudoc/image\\_30000085738601/30000085738601/RespSurv/Contents.pdf](http://www.cs.indiana.edu/sudoc/image_30000085738601/30000085738601/RespSurv/Contents.pdf)

- [Similar pages](#) - [Note this](#)

## [\[PDF\] UNSUUSEURACSB](#)

File Format: PDF/Adobe Acrobat

Santa. Rita. Agat. Yona. Apra. Harbor. Ordot. Mangilao. Asan. Hagåtña. Tiyan. Yigo.

Andersen AFB. NAS Agana. US Naval. Computer & Tele. Area Master Sta ...

[www.cs.indiana.edu/sudoc/image\\_30000053565176/](http://www.cs.indiana.edu/sudoc/image_30000053565176/)

[30000053565176/StateBasedMaps/cd108\\_GU.pdf](http://30000053565176/StateBasedMaps/cd108_GU.pdf) - [Similar pages](#) - [Note this](#)

## [Index of /sudoc/image\\_30000065775839](#)

[DIR] Parent Directory - [DIR] 30000065775839\_1/ 25-Feb-1998 05:35 - [DIR] 30000065775839\_2/ 09-Mar-1998 13:53 - [DIR] 30000065775839\_3/ 05-Mar-1998 04:58 ...

[www.cs.indiana.edu/sudoc/image\\_30000065775839/](http://www.cs.indiana.edu/sudoc/image_30000065775839/) - 2k - [Cached](#) - [Similar pages](#) - [Note this](#)

## [Index of /sudoc/image\\_30000061189928](#)

[DIR] Parent Directory - [DIR] 30000061189928\_1/ 31-Dec-1969 19:00 - [DIR] 30000061189928\_2/ 31-Dec-1969 19:00 - [DIR] 30000061189928\_3/ 31-Dec-1969 19:00 ...

[www.cs.indiana.edu/sudoc/image\\_30000061189928/](http://www.cs.indiana.edu/sudoc/image_30000061189928/) - 2k - [Cached](#) - [Similar pages](#) - [Note this](#)

## [Index of /sudoc/image\\_30000042367932/30000042367932](#)



> This comes to me following a telephone call from Alex Ragno, an  
> attorney in New Jersey, who says that there is a link from the CS  
> homepage to an old OSHA case of a client of his, Peter Regna. He would  
> like it removed from the web if possible, because I think he thinks it  
> reflects badly on his client.

>  
> What can you find out about this? (I didn't open the link, but did a  
> search on the lawyer; couldn't find anything about Peter Regno in a  
> quick search.)

>  
> Thanks.

>  
> Lucy

>  
> \*From:\* [Alexdrago124@aol.com](mailto:Alexdrago124@aol.com) [<mailto:Alexdrago124@aol.com>]  
> \*Sent:\* Tuesday, September 18, 2007 2:22 PM  
> \*To:\* [lbatters@cs.indiana.edu](mailto:lbatters@cs.indiana.edu)  
> \*Subject:\* Peter Regna

>  
> As discussed, please see link below. My client would like to know why  
> someone at IU has published this 10 year old matter on the web.

>  
> [http://www.cs.indiana.edu/sudoc/image\\_30000056020518/30000056020518\\_3/pdf\\_1996/94\\_0055.pdf](http://www.cs.indiana.edu/sudoc/image_30000056020518/30000056020518_3/pdf_1996/94_0055.pdf)

>  
> 201 207 0413.

>  
>



# SEARCHING

Indiana University Department of  
**ComputerScience**

**info**  
IU Home | IUB Ho

Accessing from ip 129.79.245.45 | Currently logged in as AnonymousUser

Search Site

Go

Home Page

Contacts

Database

Search

Help

References

RSS

Currently masking  
executable files: False  
installer files: False  
library files: False

## SUDOC Virtualization Project

This site provides access to an online database containing government documents originally published on CDROM. There are two interfaces: one enables browsing the entire collection and a Search interface (below) that allows searching for specific documents with various masks.

Welcome, new user!

fire

Search



# SEARCHING

Accessing from Ip 129.79.245.45 | Currently logged in as A

## SUDOC Virtualization Project

### 3 Result(s) for fire

0 Additional result(s) restricted based on your IP address: 129.79.245.45

<input type="text"/>	<input type="button" value="Search"/>
1. Draft environmental impact statement [electronic resource] : Monument fire recovery project and proposed non significant forest plan amendments.	30000085799520
2. EOSDIS terra data sampler #1 [computer file] : western U.S. wildfires 2000.	30000078177742
3. Multisensor fire observations [electronic resource].	30000094728650



# NEW INTERFACE

## SUDOC Virtualization Project

### **Prairie City Ranger District (Or.)**

***Draft environmental impact statement [electronic resource] : Monument fire recovery project and proposed non significant forest plan amendments.***

***Title from disc label: Monument Fire recovery project DEIS***

***Monument fire recovery project***

[Prairie City, Or.] : U.S. Dept. of Agriculture, Forest Service, Prairie City Ranger District, [2003]

1 CD-ROM : col. ill. ; 4 3/4 in.

System requirements: Computer with CD-ROM drive; Adobe Acrobat Reader.

Title from title screen.

"Malheur National Forest, Grant and Baker counties, Oregon."

"July 2003."

Shipping list no.: 2003-0031-E.

Includes bibliographical references.

"R6-MAL-03-003."

Also available via Internet from the Forest Service web site. Address as of 9/25/03:

<http://www.fs.fed.us/r6/malheur/>; current access is available via PURL.

**1.** Forest fires--Oregon--Monument Rock Wilderness. **2.** Forest fires--Oregon--Malheur National Forest.

<http://purl.access.gpo.gov/GPO/LPS36562>

1 4

Add additional metadata to this ISO (authorized users only):

**B** *I* U ABC ↶ ↷ 🔔 ☰ ☷

[Log in to add this data!](#)














[See previous edits](#)



# MIGRATION ON DEMAND

## DOC Virtualization Project

Index of image\_30000056931151/30000056931151/docs/

Migrate	Name	Last Modified
	 compress.doc	03-Mar-1990 08:03
	 format_c.doc	02-Oct-1997 05:43
	 format_e.doc	09-Oct-1997 03:09
	 format_m.doc	09-Oct-1997 02:18
	 how2zcat.doc	03-Jul-1997 03:42
-	 huswo.inv	04-Jun-1997 09:19
	 invent.doc	03-Jul-1997 03:35



# MIGRATION (MOVIE)

Accessing from ip 129.79.245.45 | Currently logged in as | Recent Changes | Logout

## SUDOC Virtualization Project

Index of image\_30000065775821/30000065775821\_6/nscat/









Migrate	Name	Last Modified	Size
	 ara96277.mpg	10-Feb-1998 10:54	1M
	 ben96299.mpg	10-Feb-1998 10:54	1M
	 ber97146.mpg	10-Feb-1998 10:54	1M
	 eat96320.mpg	10-Feb-1998 10:54	2M
	 eep97153.mpg	10-Feb-1998 10:54	3M
	 ena97010.mpg	10-Feb-1998 10:54	1M
	 enp96299.mpg	10-Feb-1998 10:55	2M
		10-Feb-1998 10:55	2M



# SPREADSHEET

## SUDOC Virtualization Project

Index of image\_30000082743968/30000082743968/volume2/c6/

Migrate	Name		Last Modified	Size
-		at06-01.pdf	19-Apr-2000 09:44	19k
		at06-01.xls	19-Apr-2000 09:44	24k
-		at06-02.pdf	19-Apr-2000 09:44	23k
		at06-02.xls	19-Apr-2000 09:44	34k
-		at06-03.pdf	19-Apr-2000 09:44	16k
		at06-03.xls	19-Apr-2000 09:44	15k



# Appendix 6-1

Appendix table 6-1.













Total, federally funded, and non-federally funded academic R&D, by basic research, applied research, and development: 1953–98  
(Percentages)

Year	Academic R&D				Federally supported academic R&D				Non-federally supported academic R&D			
	Total	Basic	Applied	Develop-	Total	Basic	Applied	Develop-	Total	Basic	Applied	Develop-
	academic	research	research	ment	academic	research	research	ment	academic	research	research	ment
		Percent of total				Percent of federally supported				Percent of non-federally supported		
1953	100.0	45.1	49.0	5.9	54.7	54.7	39.6	5.7	45.3	33.6	60.3	6.1
1954	100.0	49.0	45.3	5.6	54.7	58.7	36.2	5.2	45.3	37.4	56.4	6.2
1955	100.0	52.5	41.4	6.1	55.8	61.0	33.0	6.0	44.2	41.7	52.0	6.3
1956	100.0	56.3	37.3	6.4	56.5	64.5	29.4	6.1	43.5	45.6	47.6	6.8
1957	100.0	60.2	33.8	6.0	55.8	68.9	26.1	5.0	44.2	49.1	43.6	7.3
1958	100.0	63.5	30.9	5.6	57.0	72.1	23.4	4.5	43.0	52.1	40.8	7.1
1959	100.0	66.2	28.5	5.3	60.7	73.8	21.8	4.4	39.3	54.4	38.8	6.7
1960	100.0	68.8	26.3	4.9	64.2	75.2	20.6	4.2	35.8	57.1	36.7	6.2
1961	100.0	71.7	23.8	4.5	66.8	77.5	18.6	3.9	33.2	59.9	34.3	5.8
1962	100.0	74.0	21.0	4.0	68.0	78.5	17.0	3.0	30.0	60.1	31.0	5.0



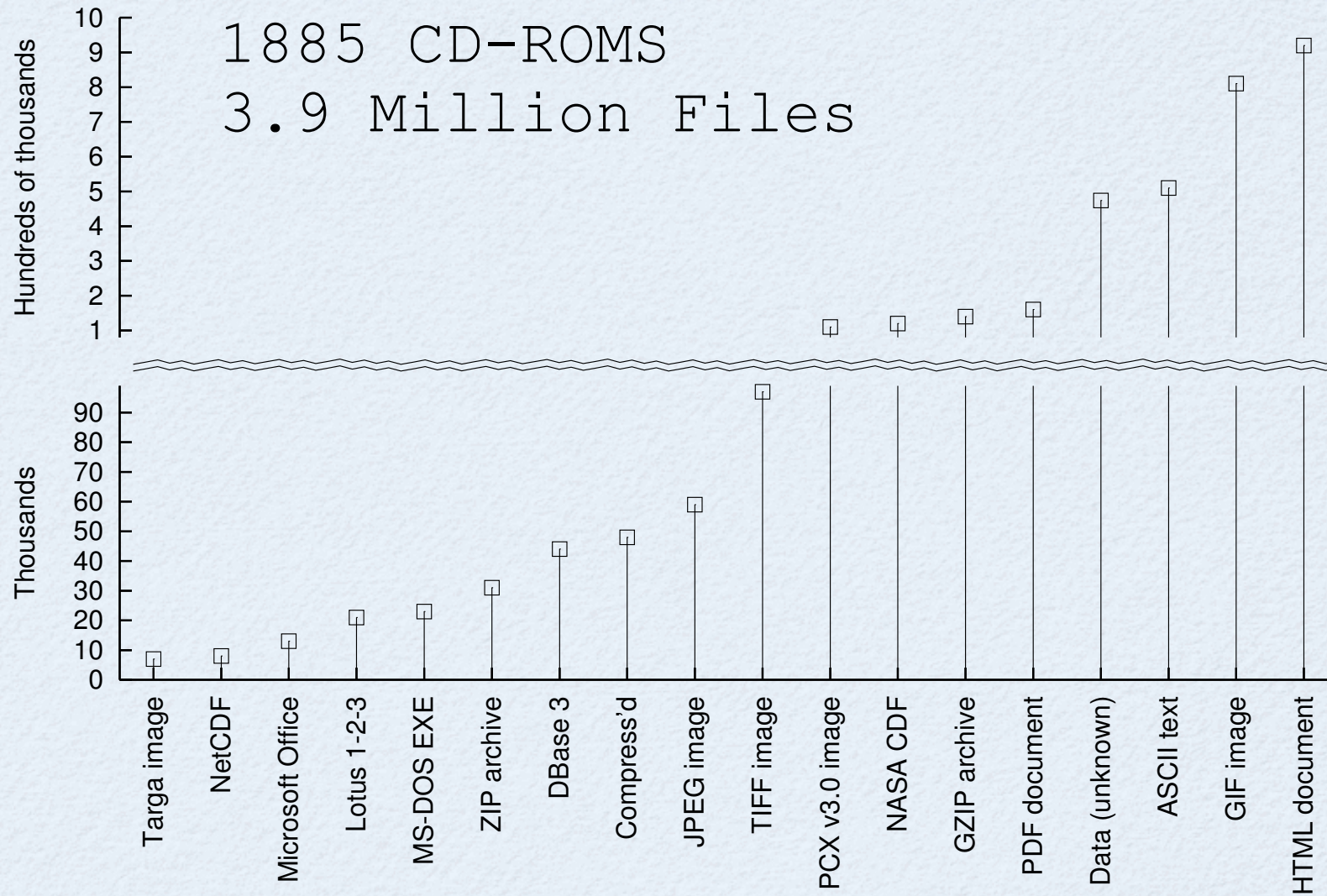
# DBASE

## Index of image\_30000035452964/30000035452964/zipcode/

Migrate	Name	Last Modified
-	 go	09-Jul-1993 03:32
	 zipnames.dbf	26-Dec-1990 10:03
-	 zipnames.ndx	27-Dec-1990 11:55
	 zip_ak.dbf	27-Dec-1990 11:06
	 zip_al.dbf	24-Oct-1990 12:28
	 zip_ar.dbf	27-Dec-1990 11:07
	 zip_az.dbf	27-Dec-1990 11:06



# LARGE EXPERIMENT





# MIGRATION CANDIDATES

	Suffix	Number in Archive	Destination
Image formats	.gif .bmp .pcx .tif .png	697624	JPG
NASA CDF	.cdf	122079	ASCII text, XML
DBase III, IV	.dbf	22851	XML / HTML
Lotus 1-2-3	.123 .wk1 .wks	19414	XML / HTML
MS Excel	.xls	12268	XML / HTML
Corel WordPerfect	.wp4(5,6) .wp .wpd	3282	Adobe PDF
MS Word 6, 95, 97	.doc .wri	1804	Adobe PDF
MS Access	.mdb	1703	XML / HTML
Video formats	.avi .swf .mov .mpg .vob	1176	Flash Video (.flv)
MS PowerPoint	.ppt	380	Adobe PDF

**Table 1.** Conversion candidates



# CONVERSION TOOLS

Image Conversion

Python Imaging Library

NASA CDF

CDF Libraries

DBase

libdbf

Excel, Lotus 1-2-3

gnumeric

Word, Wordperfect, Powerpoint

OpenOffice

Video

FFmpeg



# CONVERSION DATA

	Successful conversions	Failures	Time/File
Image formats	692118	2908	< 1s
NASA CDF	122079	3	3s
DBase III, IV	34352	718	< 1s
MS Excel	12279	11	< 1s
Lotus 1-2-3	19086	3	< 1s
Corel Wordperfect	1104	5	6s
MS Word 6, 95, 97	1267	99	7s
MS Access	2084	1	< 1s
Video formats	1176	48	6s
MS PowerPoint	161	6	5s

**Table 2.** Conversion Success



# CONVERSION RESOURCES





	Number in Archive	Original Size		Converted Size	
		Total	Mean	Total	Mean
Image formats	692118	50676MB	75KB	118074MB	175KB
NASA CDF	122079	148543MB	1247KB	9121MB	77KB
DBase III,IV	34352	65410MB	1991KB	1360MB	41KB
Lotus 1-2-3	19086	617MB	33KB	2142MB	115KB
MS Excel	12279	1038MB	6018KB	331KB	< 1KB
Corel WordPerfect	1104	179MB	167KB	227MB	212KB
MS Word 6, 95, 97	1267	405MB	330KB	196MB	160KB
MS Access	2084	12241MB	6018KB	331KB	< 1KB
Video Formats	1176	7856MB	18MB	978MB	2308KB
MS PowerPoint	161	199MB	800KB	132MB	53KB

**Table 3.** Conversion Resources



# BROWSING IN ARCHIVES






## Index of image\_30000048862662/30000048862662/aa30/unix/

Migrate	Name	Last Modified
-	 aixr.tar	07-Nov-1996 03:53
-	 aixs.tar	07-Nov-1996 03:54
-	 hpuxr.tar	07-Nov-1996 03:52
-	 hpuxs.tar	07-Nov-1996 03:53



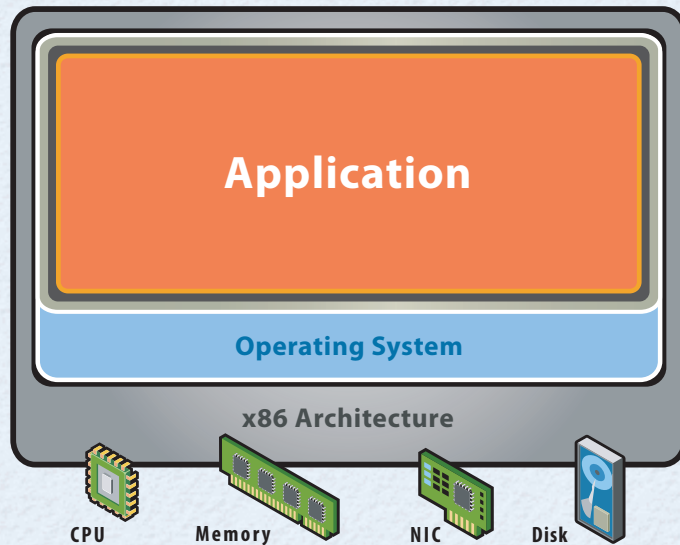
# INSIDE ARCHIVE

## Index of /mnt/isosrv/image\_30000048862662/30000048862662/aa30/unix/aixr.tar/

Migrate	Name	Last Modified
-	 aixr.tar/Reader/rs6000aix/fonts	847320804
-	 aixr.tar/Reader/rs6000aix/res	847320804
-	 aixr.tar/Reader/rs6000aix/app-defaults/AcroRead	847320788
-	 aixr.tar/Reader/rs6000aix/bin/acroread	847320787
-	 aixr.tar/Reader/rs6000aix/lib/XKeysymDB	847320805

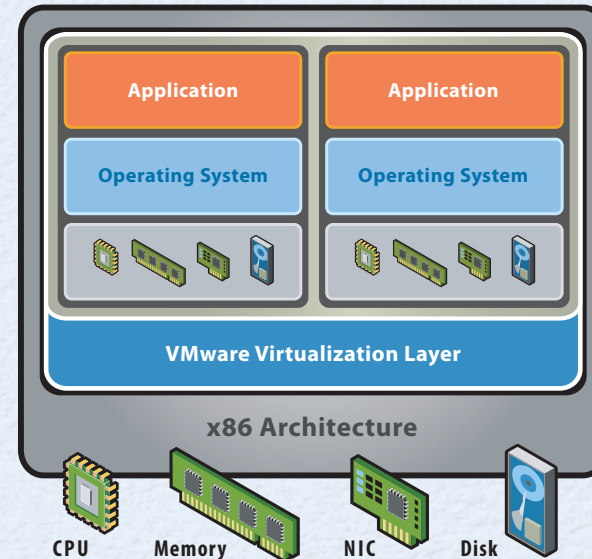


# VIRTUALIZATION



## Before Virtualization:

- Single OS image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Underutilized resources
- Inflexible and costly infrastructure



## After Virtualization:

- Hardware-independence of operating system and applications
- Virtual machines can be provisioned to any system
- Can manage OS and application as a single unit by encapsulating them into virtual machines

Source: Virtualization Overview, Copyright VMware



## Mozilla Firefox

File Edit View Go Bookmarks Tools Help

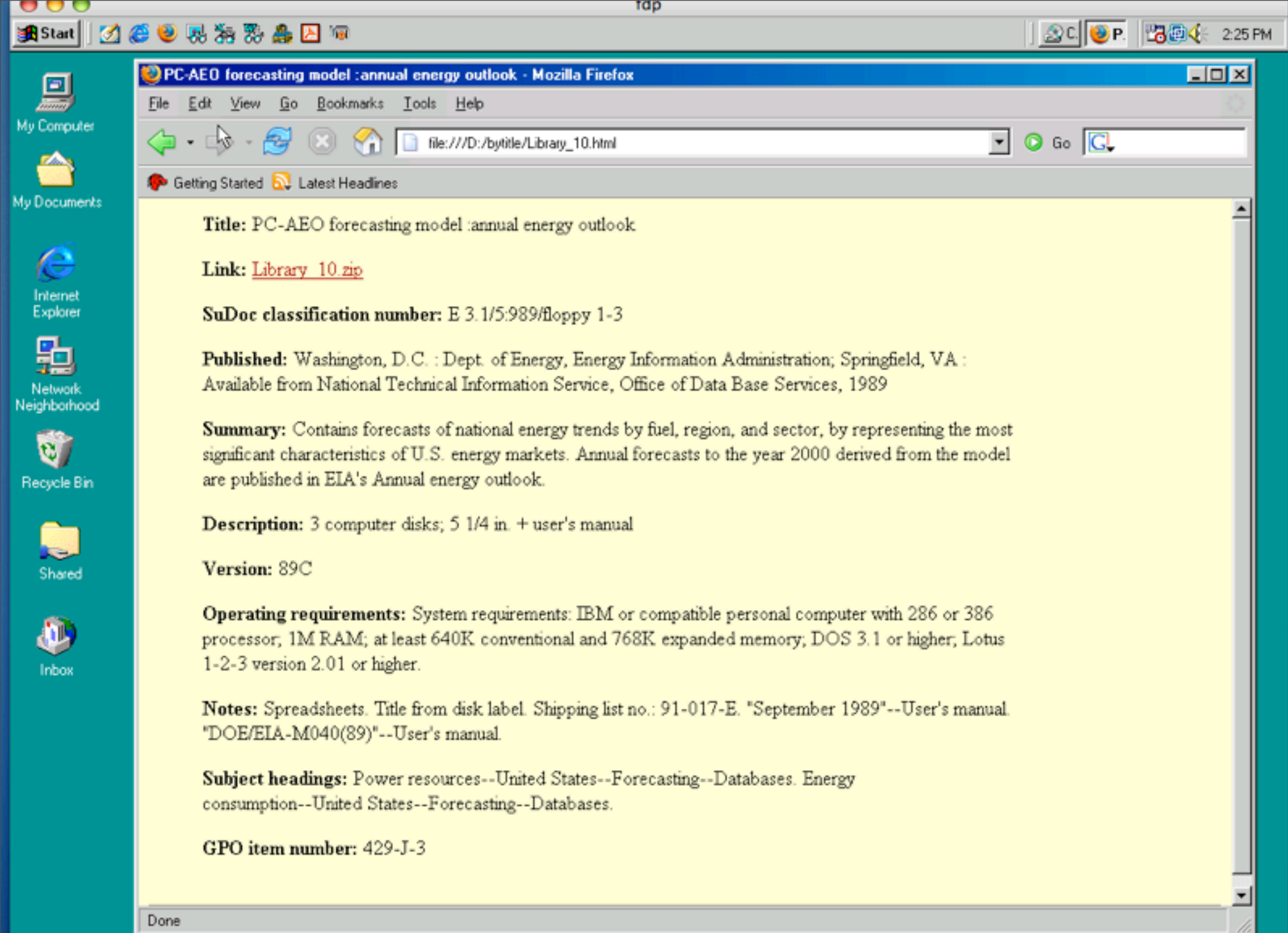
file:///D:/bytitle/bytitle.html

Getting Started Latest Headlines

- [National ambulatory medical care survey.](#)
- [National biennial RCRA hazardous waste report.](#)
- [The National health and nutrition examination surveys: a selective bibliography, 1980-93/96.](#)
- [Natural gas annual.](#)
- [NED/SIPS: a stand inventory processor and simulator program for forests of the Northeastern U States.](#)
- [NEWILD Users Manual.](#)
- [NIH/DRG/ISB.](#)
- [Non-residential buildings energy consumption survey: NBECS.](#)
- [Oak Ridge uranium market model: ORUMM.](#)
- [Oil Market Simulation Model\(OMS\).](#)
- [PCAEO 90.](#)
- [PC-AEO Forecasting Model \(Annual Energy Outlook\).](#)
- [Performance profiles of major energy producers \(selected tables\).](#)
- [A Pilot Standard National Course Classification System for Secondary Education.](#)
- [President Clinton's economic plan : and additional budgetary information.](#)
- [Public library data/National center for education statistics.](#)
- [Residential energy consumption survey.](#)
- [The role of leadership in sustaining school reform : voices from the field.](#)
- [Staff data handbook : elementary, secondary, and early childhood education.](#)
- [State energy data system census: SEDS.](#)
- [State energy price & expenditure data system : SEPEDS.](#)
- [State library agencies data FY 1996.](#)
- [Status of open recommendations.](#)
- [STF 3A, updated software.](#)
- [Student data handbook : elementary/secondary and early childhood education.](#)
- [Sugar and sweetener situation and outlook yearbook.](#)
- [Suspended-sediment budget for the Kankakee River Basin, 1993-95.](#)
- [Taxes & You.](#)
- [Toxic chemical release inventory reporting year.](#)
- [TransVU.](#)

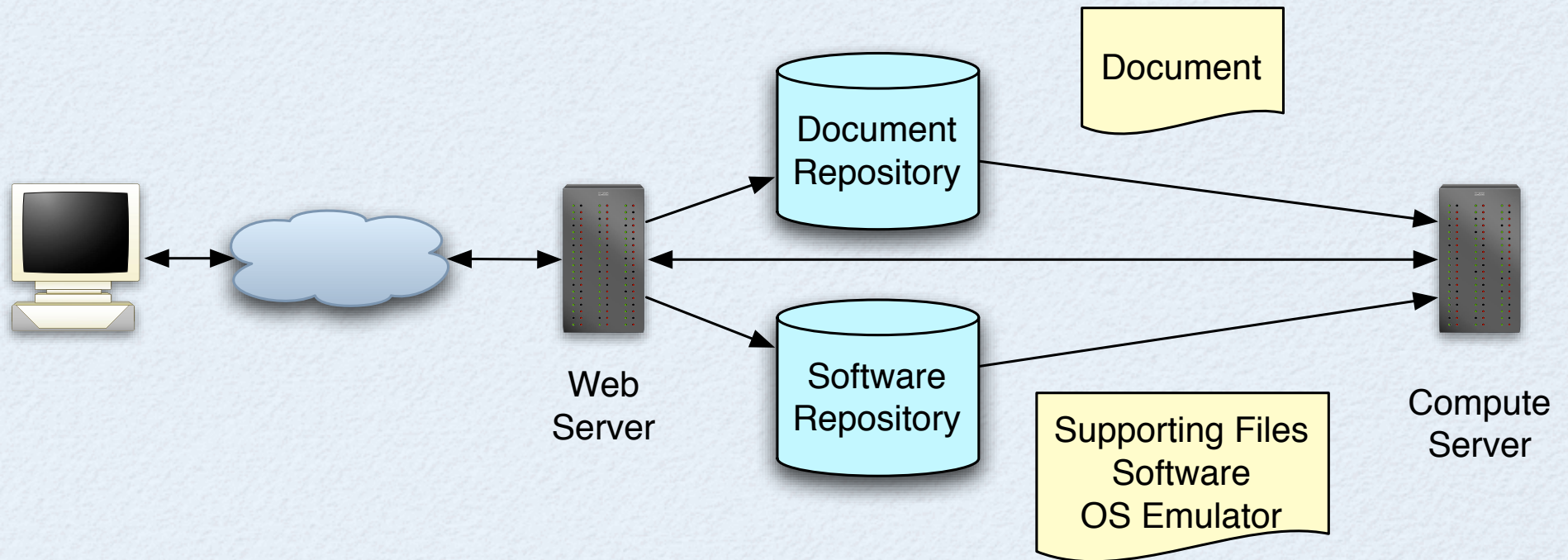
Done







# MODEL ARCHITECTURE





# ACTIONS IN RESPONSE TO PATRON REQUEST

- A pre-configured emulator is allocated
- Emulator is customized
  - Document file system mounted
  - Document specific installation executed
  - Shared file directories created for patron use
- Link to emulator and web accessible file system provided through patron browser
- Emulator executes remotely under patron control



# Indiana University Department of ComputerScience

Indiana University School of  
**informatics**  
[IU Home](#) | [IUB Home](#) | [Informatics home](#)

Iso

[Recent Changes](#) | [Login](#)

Search Site

Go

[Home Page](#)[Contacts](#)[Database](#)[Search](#)[Help](#)[RSS](#)

## Iso

### Title

World ocean atlas 1998 figures

[computer file] /

prepared by the Ocean Climate Laboratory National Oceanographic Data Center.

### Call number

C 55.297: T 24/2/DISC.1-3

### System Requirements

System requirements Microsoft: any DOS Extension that supports CD-ROM player with ISO 9660 software driver  
hard disc drive  
and VGA color monitor.

### Installation Notes

No install necessary. Data is accessible through the html file: open\_1st.html.

### Browse CD Images

[http://www.cs.indiana.edu/sudoc/image\\_30000067278824/](http://www.cs.indiana.edu/sudoc/image_30000067278824/)



Options

Clipboard

Send Ctrl-Alt-Del

Refresh



Microphone



Tools



?



Search

Bookmarks

TightVNC desktop

University School of  
**matics**  
e|Informatics home

[Recent Changes](#) | [Login](#)

http://spitz/cgi-bin/console.cgi?iso=30000067278824

**30000067278824**

- ☒ Disk 0
- ☐ Disk 1
- ☐ Disk 2

## Index of /

Name

Last Modified

Size

[Parent Directory](#)[AGR4ALL.PDF](#)

05-Oct-95 06:35

4.7M

Upload

Browse...



# HOW THIS WORK MIGHT BE USED

- Libraries share pool of software images and licenses
- Libraries share expertise in supporting various document collections
- Libraries collaborate to provide redundancy
- Patrons access from anywhere without needing to obtain or install special software



# WHY NOT JUST MIGRATION

- Loss of information -- e.g. word edits
- Loss of fidelity -- e.g. WordPerfect to Word isn't very good
- Loss of authenticity -- users of migrated document need access to original to verify authenticity
- Not always possible -- closed proprietary formats
- Not always feasible -- costs may be too high
- Emulation may necessary to enable migration



**BBC**  
**NEWS**
[OPEN](#) **BBC News in video and audio**
**News services**  
 Your news when you want it

**News Front Page**  
**World**


Africa

Americas

Asia-Pacific

**Europe**

Middle East

South Asia

**UK****England****Northern Ireland****Scotland****Wales****Business****Politics****Health****Education****Science/Nature****Technology****Entertainment**[Have Your Say](#)[Magazine](#)[In Pictures](#)

Last Updated: Monday, 2 May, 2005, 17:18 GMT 18:18 UK

[E-mail this to a friend](#)
[Printable version](#)

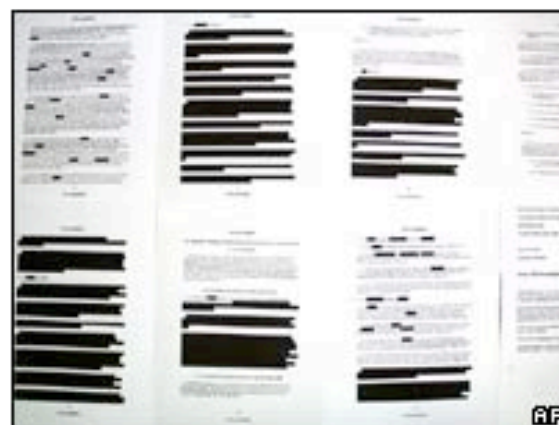
## Readers 'declassify' US document

**When news started circulating in Italy that a heavily censored Pentagon report into the death of secret agent Nicola Calipari had been decrypted, many thought it must be the work of some top-notch hacker.**

In fact, it turned out that the classified document, containing top-secret details - such as the name of the soldier who fired the deadly rounds of ammunition - could be made readable with two simple clicks of your computer mouse.

A few hours after the Pentagon published the report on its website, a few Italian readers found they could make the blacked-out paragraphs reappear by cutting and pasting them from the site into a Word document.

Salvatore Schifani, a 30-year-old IT worker, spotted the document at about 0300 local time (0100 GMT) on Saturday night.



Someone found a simple cut-and-paste job could restore the text

**BBC NEWS:VIDEO AND AUDIO**  
**How the censored parts of the report were made public**
[▶ VIDEO](#)

### THE STRUGGLE FOR IRAQ KEY STORIES

- ▶ New Iraqi plan to curb violence
- ▶ Baghdad authorities lift curfew
- ▶ Baghdad mosque attack kills 10
- ▶ Iraq war fuels terror - US report
- ▶ Saddam thrown out for third time

### BACKGROUND AND ANALYSIS



#### Tide of violence

Complete US victory in Iraq may be out of the question, argues John Simpson.

- ▶ Saddam Hussein on trial
- ▶ Iraq's federal fracas
- ▶ War justifications laid bare
- ▶ War dead figures
- ▶ Saddam trial one timeline
- ▶ Saddam trial two timeline

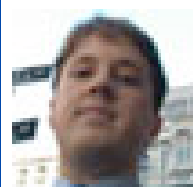


## Digital Media

# When Words Come Back From The Dead

David M. Ewalt, 12.13.05, 5:00 PM ET

### By This Author



David M. Ewalt

- [The Hartford CIO's Wireless LCD Television](#)
- [Trust](#)

NEW YORK -In the year since pharmaceutical giant **Merck** withdrew its arthritis drug Vioxx from shelves, the company has been hit with 7,000 personal injury lawsuits—one of which already cost the company \$253 million. But it could be a frequently misused feature of **Microsoft** Word that turns out to be the straw that broke Merck's back.



# STATUS

- 2300 CD-ROMs on line
- Browse only implementation (no migration)  
<http://cgi.cs.indiana.edu/~geobrown/svp>
- Migration implementation (“real soon now”)
- Public access to Emulation (??)



# ISSUES NOT DISCUSSED

- User Metadata Contributions
- Access Control
- Replication and long-term preservation
- Automated testing



# ACKNOWLEDGMENTS

- Kam Woods, Valkyrie Savage, Amanda Farag
- Lou Malcomb (IU Head GIMSS)
- Julianne Bobay (IU Head SLIS Library)

